Upper bound of typical ranks of $m \times n \times ((m-1)n-1)$ tensors over the real number field

Toshio Sumi, Toshio Sakata and Mitsuhiro Miyazaki

April 29, 2013

Abstract

Let $3 \leq m \leq n$. We study typical ranks of $m \times n \times ((m-1)n-1)$ tensors over the real number field. The number $(m-1)n-1$ is a minimal typical rank of $m \times n \times ((m-1)n-1)$ tensors over the real number field. We show that a typical rank of $m \times n \times ((m-1)n-1)$ tensors over the real number field is less than or equal to $(m-1)n$ and in particular, $m \times n \times ((m-1)n-1)$ tensors over the real number field has two typical ranks $(m-1)n-1, (m-1)n$ if $m \leq \rho(n)$, where ρ is the Hurwitz-Radon function defined as $\rho(n) = 2^b + 8c$ for nonnegative integers a, b, c such that $n = (2a+1)2^{b+4c}$ and $0 \leq b < 4$.

1 Introduction

Kolda and Bader [2] introduced many applications of tensor decomposition analysis in various fields such as signal processing, computer vision, data mining, and others. Tensor decomposition concerns with its rank and approximation of tensor decomposition concerns with typical ranks. In this paper we discuss the typical rank for 3-way arrays (3-tensors). A 3-way array

$$(a_{ijk})_{1\leq i\leq m,\ 1\leq j\leq n,\ 1\leq k\leq p}$$

with size (m,n,p) is called an $m \times n \times p$ tensor. An $m \times n \times p$ tensor of form

$$(x_i y_j z_k)_{1\leq i\leq m,\ 1\leq j\leq n,\ 1\leq k\leq p}$$

is called a rank one tensor. A rank of a tensor T, denoted by $\operatorname{rank} T$, is defined as the minimal number of rank one tensors which describe T as a sum. The rank depends on the base field.

Throughout this paper, we assume that the base field is the real number field $\mathbb{R}$. Let $\mathbb{R}^{m\times n\times p}$ be the set of $m \times n \times p$ tensors with Euclidean topology. A number r is a typical rank of $m \times n \times p$ tensors if the set of tensors with rank r contains a nonempty open semi-algebraic set of $\mathbb{R}^{m\times n\times p}$ (see [1]). We denote by $\text{typical\_rank}_{\mathbb{R}}(m,n,p)$ the set of typical ranks of $\mathbb{R}^{m\times n\times p}$. Note that

$$\text{typical\_rank}_{\mathbb{R}}(m_1,m_2,m_3) = \text{typical\_rank}_{\mathbb{R}}(m_i,m_j,m_k)$$

for any i, j, k with $\{i, j, k\} = \{1, 2, 3\}$. If s and t are typical ranks of tensors with $s \leq t$, then u is also a typical rank of tensors for any u with $s \leq u \leq t$. The minimal number of $\mathrm{typical\_rank}_{\mathbb{R}}(m, n, p)$ is equal to the generic rank $\mathrm{grank}(m, n, p)$ of the set of $m \times n \times p$ tensors over the complex number field [1]. we denote by $\mathrm{mtrank}(m, n, p)$ by the maximal typical rank of $\mathbb{R}^{m \times n \times p}$. Then

$$\mathrm{typical\_rank}_{\mathbb{R}}(m, n, p) = [\mathrm{grank}(m, n, p), \mathrm{mtrank}(m, n, p)] \cap \mathbb{Z}.$$

However, it is only known that one or two typical ranks of tensors.

For $m = 1$, the rank of $1 \times n \times p$ tensor is its matrix rank and therefore the set of typical ranks of $1 \times n \times p$ tensors consists of one number $\min(n, p)$. In the case where $m = 2$, the set of typical ranks of $2 \times n \times p$ tensors is well-known [5]:

$$\mathrm{typical\_rank}_{\mathbb{R}}(2, n, p) = \begin{cases} \{p\}, & n < p \leq 2n \\ \{2n\}, & 2n < p \\ \{p, p+1\}, & n = p \geq 2 \end{cases}$$

Suppose that $3 \leq m \leq n$. The typical rank of $\mathbb{R}^{m \times n \times p}$ is quite different from that of $\mathbb{R}^{2 \times n \times p}$. Let $\rho(n)$ be the Hurwitz-Radon number, that is, $\rho(n) = 2^b + 8c$ for nonnegative integers a, b, c such that $n = (2a+1)2^{b+4c}$ and $0 \leq b < 4$. If $p > (m-1)n$ then the set of typical ranks of $m \times n \times p$ tensors is just $\{\min(p, mn)\}$. For $p = (m-1)n$, the set of typical ranks of $m \times n \times p$ tensor is $\{p\}$ (resp. $\{p, p+1\}$) if and only if $m > \rho(n)$ (resp. $m \leq \rho(n)$).

The purpose of this paper is to give an upper bound of typical ranks of $m \times n \times ((m-1)n-1)$ tensors.

Theorem 1.1 *Let $3 \leq m \leq n$ and $p = (m-1)n-1$. A typical rank of $m \times n \times p$ tensors is less than or equal to $p+1$. In particular,* $\mathrm{typical\_rank}_{\mathbb{R}}(m, n, p)$ *is a subset of* $\{p, p+1\}$.

By [3, Theorem 1.1] we directly have the following proposition.

Proposition 1.2 *Let $m \leq n$. Suppose that $3 \leq m \leq \rho(n)$ or that both m and n are congruent to 3 modulo 4. Then*

$$\mathrm{typical\_rank}_{\mathbb{R}}(m, n, (m-1)n-1) = \{(m-1)n-1, (m-1)n\}.$$

2 A proof

In this section, we show the proof of Theorem 1.1.

First, we establish terminology.

Notation 2.1 (1) *For an $m \times n$ matrix M, we denote the $i \times j$ matrix consisting of the first i rows and the first j columns of M by $M^{\leq i}_{\leq j}$.*

(2) *For an $m \times n$ matrix M, we denote the j-th column vector of M by $M_{=j}$.*

(3) *For a square matrix M, we denote the determinant of M by $|M|$.*

(4) *For a tensor $T = (t_{ijk})_{1\le i\le m,\ 1\le j\le n,\ 1\le k\le p}$, we denote it by $(T_1;\dots;T_p)$, where $T_k = (t_{ijk})_{1\le i\le m,\ 1\le j\le n}$ for $k = 1,\dots,p$ is an $m\times n$ matrix.*

(5) *For a vector $\boldsymbol{c} = (c_1,\dots,c_n)^\top \in \mathbb{R}^n$, we denote the Euclidean norm of $\boldsymbol{c}$ by $||\boldsymbol{c}||$, that is, $||\boldsymbol{c}|| = \sqrt{\sum_{k=1}^n c_k^2}$.*

Proposition 2.2 *Let $\mathscr{R}(m,n,p;r) := \{T \in \mathbb{R}^{m\times n\times p} \mid \operatorname{rank} T \le r\}$ and τ be a canonical map from $\mathbb{R}^{m\times n\times p}$ onto $\mathbb{R}^{m\times n\times(p-1)}$ which sends $(Y_1;\dots;Y_{p-1};Y_p)$ to $(Y_1;\dots;Y_{p-1})$. If the set $\tau(\mathscr{R}(m,n,p;r))$ is a dense subset of $\mathbb{R}^{m\times n\times(p-1)}$, then $\operatorname{mtrank}(m,n,p-1)\le r$.*

Proof Since $\tau(\mathscr{R}(m,n,p;r))$ is a dense, semi-algebraic set, its interior is an open, dense semi-algebraic set and thus a Zariski open set. Therefore, $\operatorname{mtrank}(m,n,p-1)\le r$ follows. ∎

Let $3\le m\le n$, $p_0 = (m-1)n$, and $p = p_0 - 1$. We want to show that $\operatorname{mtrank}(m,n,p)\le p_0$. To do this, we show that there are a dense subset U of $\mathbb{R}^{m\times n\times p}$ and a section $s\colon U\to\mathbb{R}^{m\times n\times p_0}$ such that $\operatorname{rank} s(T)\le p_0$ for any tensor T of U. Then, by Proposition 2.2, we conclude that $\operatorname{mtrank}(m,n,p)\le p_0$.

Let W be the set consisting of $\begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}$ such that A_1 is an $(n-1)\times(n-1)$ matrix and has distinct eigenvalues, P is a nonsingular matrix so that $P^{-1}A_1P$ is a diagonal matrix, and each element of $P^{-1}A_2$ is nonzero.

Let $\mathrm{fl}_1\colon \mathbb{R}^{m_1\times m_2\times m_3}\to\mathbb{R}^{m_1m_3\times m_2}$ be a bijection defined as

$$(A_1;A_2;\dots;A_{m_3}) \mapsto \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_{m_3} \end{pmatrix}$$

and $\mathrm{fl}_2\colon \mathbb{R}^{m_1\times m_2\times m_3}\to\mathbb{R}^{m_1\times m_2m_3}$ be a bijection defined as

$$(A_1;A_2;\dots;A_{m_3}) \mapsto (A_1,A_2,\dots,A_{m_3}).$$

For $(X_1;\dots;X_m)\in\mathbb{R}^{n\times p\times m}$, we put $\begin{pmatrix} A \\ \boldsymbol{b}^\top \end{pmatrix} = \begin{pmatrix} X_1 \\ \vdots \\ X_{m-1} \end{pmatrix}$, and

$$(Z_1;\dots;Z_{m-1}) = \mathrm{fl}_2^{-1}(X_mA^{-1},\mathbf{0}) \in \mathbb{R}^{n\times p_0\times(m-1)}.$$

Now suppose that $m>3$. Let $\mathfrak{T}$ be the subset of $\mathbb{R}^{n\times p\times m}$ consisting of $(X_1;\dots;X_m)$

satisfying the following conditions:

$$|A| \neq 0. \tag{1}$$

$$|(Z_{m-1})^{\leq n-1}_{\leq n-1}| \neq 0. \tag{2}$$

$$\text{All eigenvalues of } (Z_{m-1})^{\leq n-1}_{\leq n-1} \text{ are distinct.} \tag{3}$$

$$Z_k \in W \text{ for } 1 \leq k \leq m-2. \tag{4}$$

$$\left|\sum_{k=1}^{m-2} x_k Z_k - x_m E_n\right| \text{ is irreducible.} \tag{5}$$

If $\left|\sum_{k=1}^{m-2} x_k Z_k - x_m E_n\right|$ is irreducible, then so is $\left|\sum_{k=1}^{m-1} x_k Z_k - x_m E_n\right|$ for any Z_{m-1}. Since $m > 3$, the set $\mathfrak{T}$ is a nonempty Zariski open set.

We consider the following two maps:

$$f\colon \mathfrak{V}_1 \to \mathbb{R}^{n\times n\times(m-1)}; \qquad f(Y_1;\ldots;Y_m) = Y_m(\mathrm{fl}_1(Y_1;\ldots;Y_{m-1}))^{-1},$$

$$g\colon \mathbb{R}^{n\times p\times m} \to \mathbb{R}^{n\times p_0\times m}; \quad g(X_1;\ldots;X_m) = \mathrm{fl}_1^{-1}\begin{pmatrix} A & \mathbf{0} \\ \boldsymbol{b}^\top & 1 \\ X_m & \mathbf{0}\end{pmatrix},$$

where

$$\mathfrak{V}_1 := \{(Y_1;\ldots;Y_m) \in \mathbb{R}^{n\times p_0\times m} \mid |\mathrm{fl}_1(Y_1;\ldots;Y_{m-1})| \neq 0\}.$$

Then

$$f\circ g(X_1;\ldots;X_m) = (X_m A^{-1}, \mathbf{0})$$

and more generally

$$f(\mathrm{fl}_1^{-1}\begin{pmatrix} A & \mathbf{0} \\ \boldsymbol{b}^\top & 1 \\ X_m & \boldsymbol{c}\end{pmatrix}) = ((X_m - \boldsymbol{c}\boldsymbol{b}^\top)A^{-1}, \boldsymbol{c})$$

for $(X_1;\ldots;X_m) \in g^{-1}(\mathfrak{V}_1)$.

Now, we fix $(X_1;\ldots;X_{m-1}) \in \mathfrak{T}$. Putting $(Z_1;\ldots;Z_{m-1}) = (X_m A^{-1}, \mathbf{0})$, conditions (2)–(5) hold. Since the characteristic polynomial $|Z_{m-1} - \lambda E_n|$ is divisible by λ but not by λ^2, we have $(Z_1;\ldots;Z_{m-1}) \in \mathfrak{C}$, where

$$\mathfrak{C} = \{(Y_1;\ldots;Y_m) \in \mathbb{R}^{n\times n\times(m-1)} \mid |\sum_{k=1}^{m-1} a_k Y_k - a_m E_n| < 0 \text{ for some } (a_1,\ldots,a_m)^\top \in \mathbb{R}^m\}.$$

Therefore, $f\circ g(\mathfrak{T}) \subset \mathfrak{C}$. In the previous paper [4], we showed that $\operatorname{rank} X = p_0$ for any $X \in \mathfrak{V}_1$ with $f(X) \in \mathfrak{C} \cap \mathfrak{W}_2$, where

$$\mathfrak{W}_2 := \{Z = (Z_1;\ldots;Z_{m-1}) \in \mathbb{R}^{n\times n\times(m-1)} \mid$$
$$Z_k \in W \text{ for each } 1 \leq k \leq m-1,\ |\sum_{k=1}^{m-1} x_k Z_k - x_m E_n| \text{ is irreducible.}\}.$$

For any $\boldsymbol{c} \in \mathbb{R}^n$ with sufficiently small $||\boldsymbol{c}||$ and $(Z_1;\ldots;Z_{m-1}) = ((X_m - \boldsymbol{c}\boldsymbol{b}^\top)A^{-1}, \boldsymbol{c})$, the conditions (1)–(5) hold. In addition, since $\mathfrak{C}$ is open, there exists $\boldsymbol{c}$ such that $Z_{m-1} \in W$ and $(Z_1;\ldots;Z_{m-1}) \in \mathfrak{C} \cap \mathfrak{W}_2$. Therefore, we have

$$\operatorname{rank}(X_1;\ldots;X_{m-1}) \leq \operatorname{rank} \operatorname{fl}_1^{-1}(\begin{pmatrix} A & \boldsymbol{0} \\ \boldsymbol{b}^\top & 1 \\ X_m & \boldsymbol{c} \end{pmatrix}) = p_0.$$

We complete the proof of the following theorem.

Theorem 2.3 *Let* $4 \leq m \leq n$*.*

$$\text{typical\_rank}_{\mathbb{R}}(m,n,(m-1)n-1) = \{(m-1)n-1\} \textit{ or } \{(m-1)n-1,(m-1)n\}.$$

In the case when $m = 3$, the condition (5) must be replaced as the condition that

$$\left|\sum_{k=1}^{m-1} x_k Z_k - x_m E_n\right| \text{ is irreducible.} \tag{6}$$

We show that the replacement is possible.

Theorem 2.4 *Let* $3 = m \leq n$*.*

$$\text{typical\_rank}_{\mathbb{R}}(m,n,(m-1)n-1) = \{(m-1)n-1\} \textit{ or } \{(m-1)n-1,(m-1)n\}.$$

Proof Note that the n-th column $(Z_{m-1})_{=n}$ of Z_{m-1} is zero. Then $H^{-1}Z_{m-1}H$ is equal to

$$Y_2 = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ -1 & 0 & \cdots & v_2 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & -1 & v_n \end{pmatrix} \}$$

for some nonsingular matrix H. Let P be a real vector space of dimension $n(n+3)/2-1$ with basis

$$\{x_1^a x_2^b x_3^c \mid 0 \leq a,b,c \leq n, a+b+c = n, b,c \neq n\}$$

and S be the set defined as

$$S := \{(Y_1;Y_2) \in \mathbb{R}^{n\times n\times 2} \mid Y_1 = \begin{pmatrix} u_{11} & 0 & \cdots & u_{11} \\ u_{21} & u_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ u_{n1} & \cdots & u_{n-1,1} & u_{n1} \end{pmatrix} \}$$

which is isomorphic to a vector space of dimension $n(n+3)/2-1$. Let g be a map from S to P defined as

$$g((Y_1;Y_2)) = |x_1Y_1 + x_2Y_2 + x_3E_n| - x_3^n.$$

Note that the polynomial (6) is irreducible if and only if $G((Z_1;Z_2))$ is irreducible. Now we show that the Jacobian of G is nonzero.

Suppose that for constants $c(v_j)$, $c(u_{ij})$, the linear equation

$$\sum_{j=2}^{n} c(v_j)\frac{\partial g}{\partial v_j} + \sum_{1\le j\le i\le n} c(u_{ij})\frac{\partial g}{\partial u_{ij}} = 0 \tag{7}$$

holds. We show that all of $c(v_j)$, $c(u_{ij})$ are zero by induction on n. It is easy to see that the assertion holds in the case where $n=1$. As the induction assumption, we assume that the assertion holds in the case where $n-1$ instead of n. We put

$$\lambda_j = u_{jj}x_1 + x_3 \text{ and } \mu(a,b) = \prod_{t=a}^{b}\lambda_t.$$

After a partial derivation, we put $u_{ij}=0$ $(i>j)$ and then have the following equations:

$$\begin{aligned}
\frac{\partial g}{\partial v_j} &= x_2^{n-j+1}\mu(1,j-1) && (2\le j\le n)\\
\frac{\partial g}{\partial u_{11}} &= x_1\begin{vmatrix}\lambda_2 & & & & (v_2+1)x_2\\ -x_2 & \lambda_3 & & & v_3x_2\\ & \ddots & \ddots & & \vdots\\ & & -x_2 & \lambda_{n-1} & v_{n-1}x_2\\ & & & -x_2 & \lambda_n + v_nx_2\end{vmatrix}\\
\frac{\partial g}{\partial u_{jj}} &= x_1\mu(1,j-1)\begin{vmatrix}\lambda_{j+1} & & & & v_{j+1}x_2\\ -x_2 & \lambda_{j+2} & & & v_{j+2}x_2\\ & \ddots & \ddots & & \vdots\\ & & -x_2 & \lambda_{n-1} & v_{n-1}x_2\\ & & & -x_2 & \lambda_n + v_nx_2\end{vmatrix} && (2\le j\le n)\\
\frac{\partial g}{\partial u_{ij}} &= -x_1x_2^{n-i}\mu(j+1,i-1)\begin{vmatrix}\lambda_1 & & & & u_{11}x_1\\ -x_2 & \lambda_2 & & & v_2x_2\\ & \ddots & \ddots & & \vdots\\ & & -x_2 & \lambda_{j-1} & v_{j-1}x_2\\ & & & -x_2 & v_jx_2\end{vmatrix} && (1\le j<i\le n)
\end{aligned}$$

By seeing terms divisible by λ_1 in the left hand side of (7), we have

$$\sum_{j=2}^{n} c(v_j)\frac{\partial g}{\partial v_j} + \sum_{2\le j\le i\le n} c(u_{ij})h_{ij} = 0$$

where

$$h_{ij} = -x_1x_2^{n-i}\mu(j+1,i-1)\begin{vmatrix}\lambda_1 & & & & 0\\ -x_2 & \lambda_2 & & & v_2x_2\\ & \ddots & \ddots & & \vdots\\ & & -x_2 & \lambda_{j-1} & v_{j-1}x_2\\ & & & -x_2 & v_jx_2\end{vmatrix}$$

Note that

$$
\begin{aligned}
\frac{\partial g}{\partial v_j} &= \lambda_1 \frac{\partial g'}{\partial v_j} \quad (2 \le j \le n), \text{ and} \\
\frac{\partial g}{\partial u_{ij}} &= \lambda_1 \frac{\partial g'}{\partial u_{ij}} \quad (2 \le j \le i \le n)
\end{aligned}
$$

where g' is the determinant of the $(n-1)\times(n-1)$ matrix obtained from $x_1Y_1+x_2Y_2+x_3E_n$ by removing the first row and the first column minus x_3^{n-1}. Therefore we have

$$c(v_j) = c(u_{ij}) = 0 \quad (2 \le j \le i \le n)$$

since $\frac{\partial g'}{\partial v_j}, \frac{\partial g'}{\partial u_{ij}}$ $(2 \le j \le i \le n)$ are linearly independent by [4, Lemma 5.2]. By (7), we have

$$c(u_{11})\frac{\partial g}{\partial u_{11}} - \sum_{i=2}^{n} c(u_{i1})u_{11}x_1^2x_2^{n-i}\mu(2,i-1) = 0. \tag{8}$$

By expanding at the n-th column, we have

$$\frac{\partial g}{\partial u_{11}} = (v_2+1)x_1x_2^{n-1} + \sum_{i=3}^{n-1} v_ix_1x_2^{n-i+1}\mu(2,i-1) + x_1(\lambda_n+v_nx_2)\mu(2,n-1)$$

and then the equation (8) implies that

$$
\begin{aligned}
(c(u_{11})(v_2+1)x_2 - c(u_{21})x_1)x_1x_2^{n-2} + \sum_{i=3}^{n}(c(u_{11})v_ix_2 - c(u_{i1})x_1)x_1x_2^{n-i}\mu(2,i-1) \\
+ (c(u_{11})(\lambda_n+v_nx_2) - c(u_{n1})u_{11}x_1)x_1\mu(2,n-1) = 0.
\end{aligned}
$$

By seeing the coefficient of $x_1x_2^{n-1}$, we have $c(u_{11}) = 0$. Further, by seeing the coefficients corresponding to x_3^s, $0 \le s \le n-2$ in the equation (8), we have $c(u_{i1}) = 0$ for $2 \le i \le n$. Therefore, we conclude that $\frac{\partial g}{\partial v_j}, \frac{\partial g}{\partial u_{ij}}$ are linearly independent, which means that the Jacobian of g is nonzero.

Therefore the set of $(Z_1,\dots,Z_{m-1}) \in \mathbb{R}^{n\times p}$ such that $(Z_{m-1})_{=n} = \mathbf{0}$ and the polynomial (6) is irreducible is a Zariski open set. Hence, the condition (5) is replaced with (6). ∎

References

[1] S. Friedland. On the generic and typical ranks of 3-tensors. *Linear Algebra Appl.*, 436(3):478–497, 2012.

[2] T. G. Kolda and B. W. Bader. Tensor decompositions and applications. *SIAM Review*, 51(3):455–500, September 2009.

[3] M. Miyazaki, T. Sumi, and T. Sakata. Typical ranks of certain 3-tensors and absolutely full column rank tensors. *preprint*, Dec. 2012.

[4] T. Sumi, T. Sakata, and M. Miyazaki. Typical rank of $m \times n \times (m-1)n$ tensors with $3 \leq m \leq n$ over the real number field. *preprint*, Dec. 2012.

[5] J. M. F. ten Berge and H. A. L. Kiers. Simplicity of core arrays in three-way principal component analysis and the typical rank of $p \times q \times 2$ arrays. *Linear Algebra Appl.*, 294(1-3):169–179, 1999.